DOCUMENT RESUME

ED 260 166 UD 024 386

AUTHOR Marklund, Sixten
TITLE From Assessment of Individual Students to Evaluation of School Systems.
PUB DATE 17 Dec 84
NOTE 20p.
PUB TYPE Reports - General (140)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Achievement Tests; *Comparative Analysis; *Educational Assessment; Educational Policy; Elementary Secondary Education; Evaluation Criteria; Foreign Countries; Outcomes of Education; *Program Evaluation; School Districts; *School Effectiveness; *Standardized Tests; *Test Use
IDENTIFIERS *Sweden

ABSTRACT

Individual standardized test scores in major school subjects can be combined to evaluate school programs and systems. Recently, Sweden experimented with such a practice. Test scores in different subjects (Swedish, English, mathematics) were gathered, units of different size (schools, municipalities, regions) were compared, and inter-school variances, inter-municipality variances, and inter-regional variances were determined. Problems with the use of individual assessment as a foundation for system evaluation include the fact that standardized test use in Sweden is voluntary and teachers, fearing that student results might be used to evaluate them as well as their students, might not use the tests. And if their use is made compulsory it might steer teaching and learning in an undesirable direction. Nevertheless, this method does have its uses. Taking the evaluation process a step further, data from an international evaluation study by the Association for the Evaluation of Educational Achievement showed that the between-school variance was lower in Sweden than in some other comparable countries. "Good" and "bad" student results were as frequent in Sweden as in the other countries, but were not as concentrated in special "good" schools and "bad" schools. (KH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Sixten Marklund
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED260166

Sixten Marklund
Institute of International Education, University of Stockholm, Sweden.

FROM ASSESSMENT OF INDIVIDUAL STUDENTS TO EVALUATION OF SCHOOL SYSTEMS

Abstract

This paper holds a discussion of a need for evaluation of educational programmes and systems as well as of individual results. It is argued that individual achievements in central school subjects assessed through standardized tests can be used for the evaluation of school programmes and systems. An example of how this was done in Sweden is given. Tests in different subjects were combined (Swedish, English, Mathematics). Units of different size (schools, municipalities, regions) were compared. The results in these units were not restricted to level of assessment but also concerned the variation between the studied units. Inter-school variances, inter-municipality variances and inter-regional variances were determined. A problem with the use of individual assessment as a foundation for system evaluation is that teachers sometimes hesitate to report the results of the standardized achievement tests in their classes as long as these results are given through the teachers' use of instruments that are voluntary and supplied as aids for their own internal evaluation. If the use of these instruments is made compulsory it might steer teaching and learning in an undesirable way.
- When using data from the IEA project it was found that the between-school variance was lower in Sweden than in some other comparable countries. "Good" and "bad" student results were as frequent in Sweden as in the other countries, but these were not to the same degree as in these other countries concentrated to special "good" schools and "bad" schools.

KEYWORDS (IBE): Programme evaluation, test results, testing programmes, regional surveys.

UD 024386

1 THE NEED FOR EVALUATION OF EDUCATIONAL PROGRAMMES AND SYSTEMS

1.1 Individual Results as Indicators of School Results

Assessments of outcomes of education usually focus on results of single individuals in single grades, subjects or courses. Test results tell whether the individual has passed or failed, what grade he has achieved and how he stands in relation to other individuals in the population concerned.

This kind of assessment and evaluation in education has a long history and is well developed. As concerns the evaluation of educational programmes and systems the situation is the reverse. This kind of evaluation is new and fairly undeveloped. An educational programme concerns many students, subjects and courses. A programme can be training for a trade, for a profession or for a restricted examination. Such programmes usually also have specialized evaluation criteria to be applied, not only on individual students but also on students as a group.

An educational system is an even wider concept. It can be a whole school (i.e. primary school) for a whole nation, state or region.

Assessment of individual results is one end of a dimension, where evaluation of school systems makes up the other end. Between these two poles there are many intermediate and mixed types of evaluation. Individual assessments can be combined and structured in different ways to form the basis of evaluation of more complex and comprehensive programmes or systems. It is becoming increasingly evident that such complex and comprehensive evaluations, to be of value for the improvement of programmes and systems, require a more systematic approach. Also new instruments, techniques and procedures are often needed. The attempts to develop such evaluation procedures have so

far been only piecemeal. The reasons for this are many, one being the simple but overwhelming fact that this kind of evaluation is strongly dependent on educational policy and educational values, not only on the institutional level but also on the local governmental and national governmental level. Realizing this we too easily forget that the assessment of individual results also has strong political implications.

Nevertheless it seems obvious that evaluating the effectiveness of programmes and systems is becoming urgent. Accountability is asked for. Most countries are short of money for education. Not only policy-makers and administrators of education but also purchasers of education are interested to see if they are successful or not in their ambitions to create a good education. Taxpayers want to know what they are offered. Besides: education needs a better adaption to modern society and modern industry, it is said. Students claim the right to participate in decisions on their education, etc.

Governments and responsible authorities have, during the post-war period, also taken a number of steps to improve education and adapt it to new individual and societal demands. For instance through a) extending the period of compulsory training at school (in industrialized countries usually to the age of 15 or 16), b) broadening secondary education to include a large part of each year cohort of students (from elite education to mass education), c) co-ordination of general-academic and technical-vocational education into new programmes and d) comprehensivization of the school structure with different programmes within the same organization. In some systems the organizational diversification of students into different fields of study has been postponed, final examinations have been changed, access to higher education has become easier, etc. According to some studies and observations the standard of student knowledge is on the decline. Other problems, such as the lack of "basic" skills and the decline of student interest for science, are also mentioned.

Changes of this kind (the list could be extended and specified) articulate the need for an evaluation which goes beyond an assessment of single individuals in single subjects. At least two types of more complex evaluation are needed. One type can be called a "total curri-

culum" evaluation. Knowledge, skills and attitudes of the individual student, seen as a unit, here make up the subject of evaluation. A second type is the result (product) of a total group of students in one or more evaluation criteria. In this second type of evaluation differences between individuals, schools and regions become important subjects of analysis.

1.2 Evaluation Criteria, Instruments and Procedures

The great amount of evaluation criteria and the complexity of these criteria immediately appear as a first and somewhat discouraging difficulty. Effectiveness of a programme or a system must be related to its goals, aims or objectives. The goals of educational systems (i.e. compulsory school in country X) are usually only vaguely defined in laws and statutes.

The goals of institutions within a system are usually more precisely stated through the structure, curriculum and time-table of the institution, through the textbooks used etc, but the institutions - as well as the systems - also have goals, which are not explicitly recognized. The extraction from this pile of external and internal conditions of assessable criteria is a difficult and sometimes also delicate task. But it is certainly not impossible to make a list of evaluation criteria (it has been made in a number of restricted studies) and also to bring these criteria into some structured order. A programme or system evaluation project has to start with a throroughgoing study of the evaluation criteria. Such a study should embody opinions from different kinds of teaching-learning experts, policy-makers and administrators of the system as well as evaluation experts. It is especially important that those who once planned the organizational and economic frames of the system participate in the definition of evaluation criteria.

From criteria one should proceed to instruments. Assessment of individual results (knowledge, skills, attitudes) quite naturally are necessary elements here, but there are also instruments of other kinds for assessing the institution, programme or system as a unit. Here it is important to find out what kinds of instruments that exist

and how they have developed in recent years.

The same questions could be asked about the evaluation procedures and their relationships to evaluation criteria and instruments. Evaluations of institutions are often made through inspections, accreditations and auditors' revisions, but are seldom done in a scientifically solid way. The fact that effectiveness of a system comprises policy questions and value questions does not put it out of reach of evaluation. But goals, instruments and procedures should be discussed and specified. Examples of some important questions are: On what administrative level of an education system should evaluation be executed (local, regional, national)? How should an evaluation team be composed? How often should a system be reviewed? How should system evaluations be reported and published?

A simplified design of a study in order to define and specify evaluation criteria is given in the following two-dimensional matrix.

Levels of evaluation	Goals/ Criteria	Instruments	Procedures
Institutional	A	B	C
System	D	E	F

Each one of the six cells A-F should be studied. The studies can be put together A+D, B+E and C+F. Another model of study could be analyses of A+B+C seen as a unit, and D+E+F treated the same way.

A key question is here, whether a system (i.e. compulsory school in a region) can be seen only as a sum of its subsystems or institutions (schools), thus whether it is sufficient in a system evaluation to analyse and evaluate its parts. Some evaluators argue, that the totality here is more than a sum of its parts. This might be true, but my point here is that a system evaluation according to a "summing-up" design is a good starting point and certainly better than no system evaluation at all.

1.3 Some Methodological Questions

The evaluation of programmes and systems rests on the basic assumption that "effectiveness" means something good for the single student. The optimal development of the individual is the ultimate goal also for institutions and systems. This can be a highly policy-oriented question. Effectiveness for whom? For what?

In a big international evaluation study by "The Association for the Evaluation of Educational Achievement", abbreviated IEA it was found that all participating countries (school systems) in their laws about education (sometimes even as parts of the nation's constitution), identify "equal opportunity" as one of the overriding educational goals. It was obvious, however, that different countries (systems) here arrived at different operational definitions of "equal opportunity" and therefore also ended up with partly different evaluation criteria. But even here it was possible to find common cores.

The following discussion is based on the assumption that effectiveness of systems can be determined through the assessment of individual students, provided that the instruments stand for measurements of major elements of the system.

Two types of scores of assessments are here of utmost importance: 1) the mean score, in whatever criterion or type of assessment defined, of the institution and the system; 2) the variability (standard deviation or other kinds of corresponding estimation) of the same assessment. The first one indicates the level of the institution or the system. The second one tells how many students who fall under an acceptable minimum standard (and therefore might need extra support) and how many who are extra high (and therefore perhaps should be observed and treated in a special way). Quite naturally it is here important to have not only traditional achievement tests scores but also estimations of different kinds of non-cognitive results of education.

Considerations of this kind, emerging from the egalitarian goals, underline the need for different analyses of varance, such as:

inter-individual variance
intra-institutional variance (classes, groups)
inter-institutional variance (schools)
intra-system variance (regions, districts)

This could help teachers, heads of schools, administrators and policy-makers to find out, where a system or an institution has its strong points and its weak points and how available resources therefore might be used in other ways to improve the results.

Such analyses of variance are easily done with assessments according to standardized achievement tests. But in principle they could and should be done with criteria like kind and degree of students participation and decision-making within institutions and systems, how the training within institutions and systems is related to changes in the society, industry, labour market etc. Obviously, "effectiveness" can also be defined in many, often quite pragmatic ways, as for instance the pass-fail frequency, frequency of class-repeating, annual cost per student, total public experditure per graduated student (per certificate).

2 SOME EXPERIENCES IN FIELD-WORK ON EVALUATION THROUGH ACHIEVEMENT TESTS

Realizing the growing need for instruments for system evaluation and programme evaluation we ought to start by asking if and how the already existing instruments for individual assessment can be used for this purpose. Here I will give some examples of how this has been practiced in the Swedish school system in recent years. Hopefully, the experiences from this have some general interest. It is necessary here, however, to start with a brief description of the Swedish system or assessment and use of evaluation instruments.

The Swedish school system is unitary in the sense that the same general and specific aims are pursued in the same kind of educational institutions all over the country. Thus all those studying any given subject at the same level usually follow the same curriculum and have the same number of weekly periods.

The ideology underlying the structure of the school system is the principle that all students - irrespective of social stratum or geographical position - are to have the same opportunity of developing their personalities and acquiring knowledge and skills as far as their individual aptitudes and abilities will allow.

Sweden has adopted a "rolling reform" in education, i.e. the educational system is revised continually. On the basis of experiences in school and special investigations the Ministry of Education proposes bills to Parliament outlining changes of major principles concerning school organization and general aims. The ensuing act is implemented by the second highest school authority, the National Board of Education (abbreviated NBE in this paper) which provides every school in the country with information on modifications and innovations decided on by Parliament, supplemented by more detailed guidelines and recommendations worked out by the NBE. This information is contained in a handbook - "läroplan" (referred to as the Curriculum in this paper) stating the over-all aims of education as well as the aims and objectives of all subjects being taught, outlining the syllabus and giving the guidelines for each subject, and discussing teaching methods and materials. There is one such Curriculum for the compulsory school and one for the Upper Secondary School. It should be noted that although guiding principles for the teaching of different subjects are laid down, it is clearly stated that within the general framework it is ultimately up to the teacher to develop the type of approach that best suits his or her personality. Single schools or groups of schools also run experimental programmes.

Certain general aims are to be pursued, regardless of what subject is being taught. Among the general aims that are to be emphasized throughout school are clarity and order of thought, the ability to think critically and independently, the ability to resist tendentious influence, the ability to analyse and deal rationally with a problem as well as the ability and willingness to co-operate with others.

Specific aims have also been set for each subject. They are mostly stated in rather general terms, but the Curriculum also indicates well-defined study areas and activities that are to be included in the learning process. The evaluation of skills and

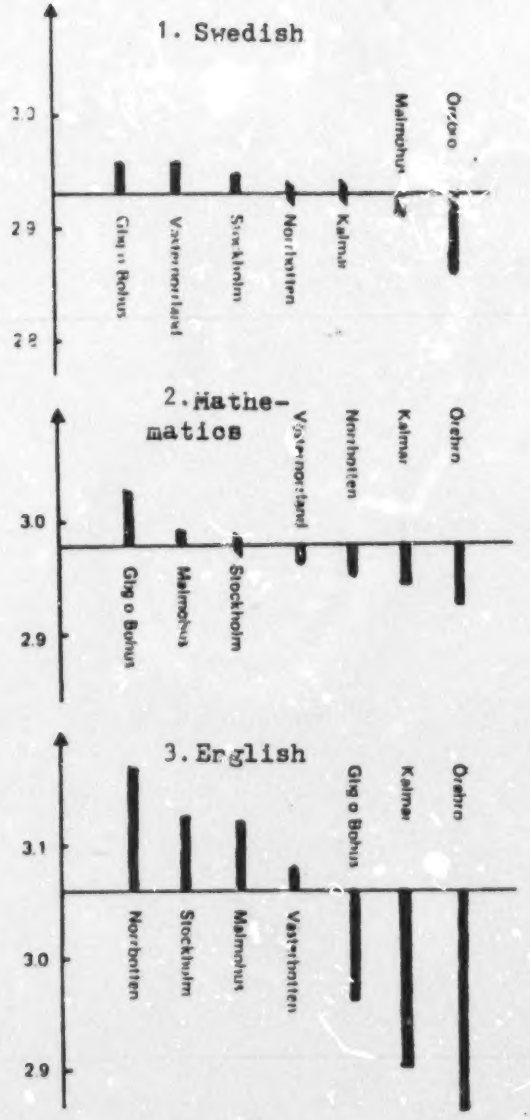
knowledge has to be based on an analysis and interpretation of the aims laid down in the Curriculum.

The outcome of such an analysis may vary owing to the subjective judgement of the person or persons making it. Consequently, the teaching of the same subject may differ, to a certain extent, according to teachers' opinions on what are the most important aspects. It is also essential that there is room for variation and change in order to adapt class-room work to the students' individual interests and aptitudes. Naturally, the freedom thus granted to teachers and students is limited by the necessity of conducting work in such a way that the stated aims can be attained. When it comes to large-scale assessment of achievements in any given subject, which requires the construction of special test batteries, it is imperative to reduce the subjective element in the interpretation of aims to a minimum. For this reason, decisions on the nature and extent of such evaluation are always based on the consensus of a fairly large group of experts.

3 REGIONAL AND LOCAL DIFFERENCES IN STUDENT ACHIEVEMENT

The way student achievements are assessed in Swedish schools by using standardized and non-standardized tests has here been surveyed rather extensively in order to illustrate the possibilities of using individual results as a basis for the evaluation of the school system as a whole.

Statistics of student marks in single subjects and single grades for the country as a whole have occasionally been collected and analysed by NBE. The reason for this has mainly been to see, if the marking system functions satisfactorily to guarantee that marks from different schools are comparable. But school marks have in recent years not been used as a basis for analysis of variance of school standards in different districts and regions. Here the standardized tests have been seen as a better instrument, although they - as has been stated above - cover only restricted parts of the outcomes.



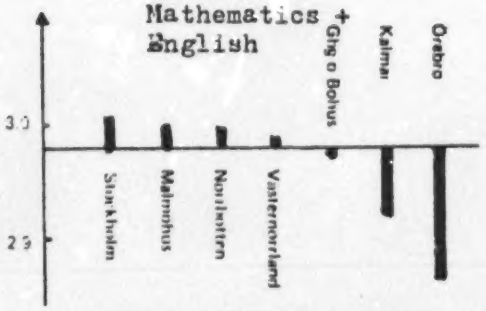


Diagram 1-4.

Mean achievement in seven counties in Swedish (1), Mathematics (2), English (3) and the three subjects together (4). Grade 6.

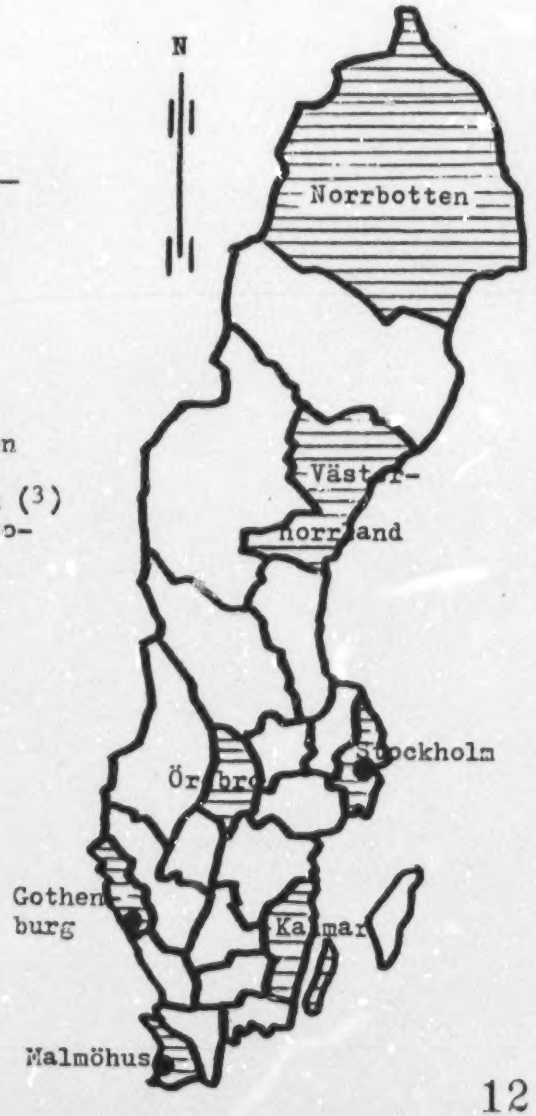
SWEDEN has

24 counties ("län")

278 municipalities

3000 schools (rector districts)

8.2 million inhabitants



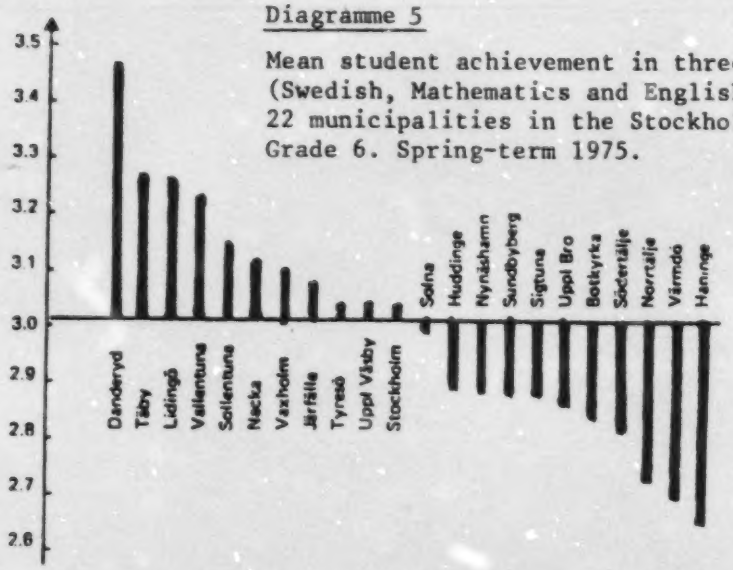
Such a study was done in 1975 by the NBE. Standardized achievement tests were applied to students in grade 6, the end of primary education (student age 12), in seven of the altogether 24 Swedish counties. Three test batteries were used: Swedish, Mathematics and English. Evaluation was made in a five point scale with 3 as the national mean and 1 as the unit of standard deviation.

A reason to study the variance of student results in units of different size was that this variance was seen as an indicator of equality of educational opportunities. According to the egalitarian goal, the difference in assessment between schools, between municipalities and between regions within the nation should be reduced or even smoothed out.

Diagrams 1-4 say that the seven counties are different in mean student achievement in Swedish, Mathematics and English (Diagrams Nos. 1-3 and in mean achievement of these three subjects taken together (Diagram 5). As can be seen there are certain regional differences in mean achievements.

Diagramme 5

Mean student achievement in three subjects (Swedish, Mathematics and English) in the 22 municipalities in the Stockholm County. Grade 6. Spring-term 1975.



These results were presented by the NBE in an official report to the Swedish Government and to members of the Swedish "Riksdag" (Parliament). The results caused much discussion in Sweden as why such differences exist, what correlates they have in other regional

variations and what one could and should do to equalize educational opportunities. The analyses of data went on to compare different municipalities within each county. Diagram 5 shows how the 22 municipalities in the Stockholm County vary in student achievement. The difference between the top (Danderyd, 3,48) and the bottom (Haninge, 2,64) is considerable. The largest municipality, the City of Stockholm, is, as we can see from the diagram, on the middle of the range (3.0).

The variance between municipalities (diagram 5) is as expected greater than the variance between counties (diagrams 1-4), just as the between school variance is greater than either that for counties or municipalities. This rule, which says that the variance between units increases when the size of the unit decreases, can be studied in table 1.

Table 1. Standard deviation of achievement in three subjects (Swedish, Mathematics and English) in grade 6.

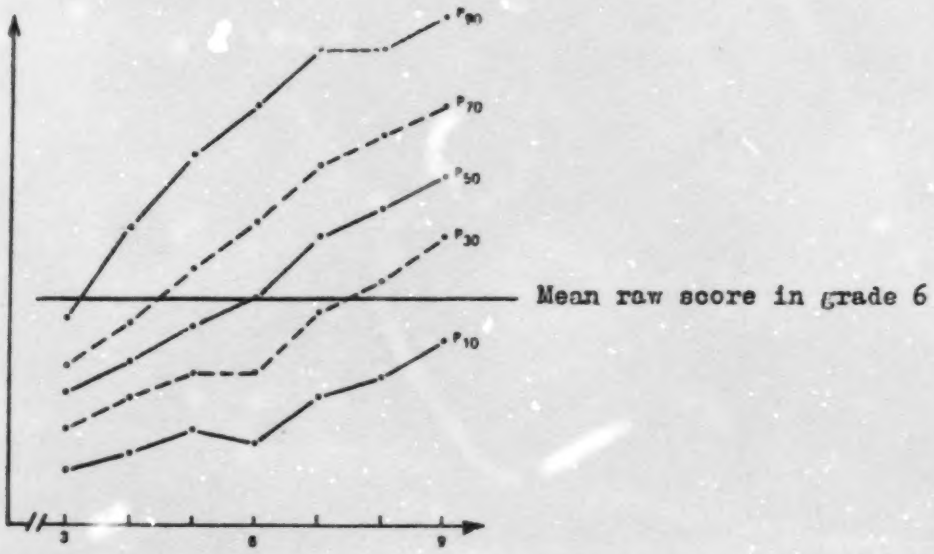
Subject	Standard deviation between			
	Counties	Municipalities	Schools	Students
Swedish	0.04	0.17	0.28	1.03
Mathematics	0.12	0.24	0.36	1.05
English	0.04	0.18	0.31	1.02
Swedish + Mathematics + English	0.06	0.18	0.30	0.91

Studies of this kind have shown that high-achieving students and low-achieving students are found in all schools and in nearly all classes. But the proportion of these two groups vary. Special concern has been given to the low-achieving group. In one of the municipalities (Linköping, 106.000 inhabitants) a study was made about student achievement in reading and writing at the end of compulsory

education, i.e. grade 9. The results of the students in grade 9 are depicted in diagram 6, specified for percenti es 10, 30, 50, 70 and 90. As can be seen in the diagram 15 percent of the students had a reading score which was lower than the mean score in grade 6. This level, mean level of grade 6, was here defined as the minimum level of competence for compulsory school leavers. The 15 percent of students with results lower than this minimum were classified as "functionally illiterate".

In the public discussion that followed the school was criticized as responsible for creating functionally illiterate citizens every year. On the other hand some people felt that the mean level of achievement for grade 6 students was too high a minimum to set. This and corresponding results from system evaluations has begun to create an increasing awareness of how the school works. They have made people start thinking of how available resources might be better used. A law enacted by Parliament in 1977 now gives the local education authorities increased possibilities of use state grants the way they decide themselves in that 25 percent of state grants are not ear-marked in advance by state authorities. Spending of this "free quarter" forces system evaluations on a broader scale on the municipal level as well as regional and national level.

Diagram 6. Increase in reading comprehension from grade 3 to grade 9. (Grundin 1975)



4 HOMOGENEITY OF NATIONAL SYSTEMS

A study of the kind that NBE did (preceding section of this paper) only tells if and how different regions, municipalities, schools and classes vary. It can prove that one region is "better" than another, but it says nothing of the "Swedish standard". Is the school system as a whole "good" or "bad"?

In the international evaluation study mentioned above (IEA) a large number of such comparisons between the participating countries were made. Student attitudes to school and education as well as student achievements in school subjects were compared. In these comparisons the interest mainly focussed on the mean level of each country. So in a way this became a kind of educational "olympic game", although the international project team denied that this was the meaning. For instance, in many subjects Japan proved to have the highest mean standard, which quite naturally caused much discussion and questions. Considerably less interest was attached to the comparisons between the nations of the intra-national variance. Table 2 shows the results of one of these comparisons, namely the variation of schools in literature knowledge in five countries, England, Finland, Italy, Sweden and the U.S.A. for 14 year old students. The differences between "good" schools and "poor" schools varied in these countries. As can be seen all five countries have both high-achieving and low-achieving students (middle column, indicating the between-student variance). The left column is an estimation of how schools vary in mean results. High figures here tell that there is a great difference between "good" and "poor" schools. Here Sweden proves to have the lowest between-school variance. Swedish schools thus here are mutually more alike than schools in the other four countries. In the right column we can see how the between-school variance is related to the total variance. Here the Swedish figure is the lowest (15%), which means that Sweden is the most homogeneous country in this respect.

Table 2. Literature knowledge according to IEA, 14-year-old students (Purves, AC, Literature Education in Ten Countries, Stockholm 1973, p 292, 294).

Country	Standard deviation		Between-school variance in percent of total variance
	schools	students	
England	2.1	4.1	26
Finland	1.7	3.4	25
Italy	1.8	3.2	31
Sweden	1.5	3.8	15
USA	2.1	4.2	25

It can certainly be discussed whether this is good or bad for Sweden. Official Swedish policy says that students should have the same opportunity to get a good education, regardless of socio-economic status and geographical location. In respect of this goal the result must be seen as good. Sweden has roughly the same proportion of "good" and "bad" students as the other four countries, but these are not concentrated in specific schools to as great an extent as in the other four countries.

5 INTERNAL OR EXTERNAL EVALUATION

The method of making system evaluation by means of combining individual assessments at the group level or even at the total population level gives rise to a series of problems, which are not only important but also difficult to solve. One of these is the question if a system evaluation should be an external or an internal one. Evaluation instruments (tests) can be constructed and applied by evaluators within as well as outside of a system. The important point here is noth whether instruments are constructed or applied by insiders or outsiders but if those who interprete and value the results of these testings are from outside or inside of the system.

The IEA type of evaluation was quite clearly an external evaluation. NBE's evaluation of student achievement in different regions, municipalities and schools was also a kind of external evaluation but this evaluation was made as a composition of internal evaluations. Teachers have applied the instruments in their own classes, added the points for single students in single subjects, put these together and sent the results to NBE. The fact that NBE's external evaluation makes up an analysis of single teachers' and schools' internal evaluation might afford difficulties. Three of these are the following:

1) The use of standardized achievement tests of the kind we described above is not compulsory for the teachers. In the Basic School (grades 1-9) these tests are fully voluntary for the teacher, he can take them or leave them. The reason for him to use them is predominantly to find out the standard of his class (mean and standard deviation) and to use this information in his own marking of single students. If the teacher is asked to report his students' results to NBE he might find himself evaluated as a teacher and therefore abolish the use of these voluntary standardized tests.

In the post-compulsory schools (grades 10-12) it is obligatory to have these tests, but the teacher is here free to make his own tests (after agreement by his headmaster) and use these instead of the standardized central ones. NBE can have the teacher send in his results, but these are normally not fit for use in an external evaluation.

2) If the central standardized tests are made compulsory in grades where they so far are not, which of course is possible to do, we have to face the fact that these tests are strongly steering for the contents and methods of teaching and learning. In the long run this could become a hinderance for the development of education. Protests against the use of central standardized tests in grades 10-12 have - all since they were first introduced in 1967 - occurred among groups of students. On the whole the system is still accepted and seems to work satisfactorily but there is a trend among teachers to replace the standardized tests by teacher-made ones. Quite naturally this will increase the freedom of teaching and learning, but at the same time it will reduce the comparability of indi-

vidual marks in different classes and schools. In a school system where individual marks make up the sole basis for selection for further studies there is also a great risk for inflation of marks.

3) If the test results are used as one of the foundations for individual marking (in Sweden often the most important foundation), the technical characteristics and qualities of the marking system become decisive also for the testing system. The so-called relative marking system (2.2 above) has been discussed and criticized all since it was first introduced in 1962. This system, it is said, is good for ranking of students but not for evaluation of actual results. Defenders of it argue, that it works the same way as the previous system of so-called absolute marking, only that it improves the comparability of marks when a selection of students has to be made. It seems obvious, however, that the strong linking of these three things together, the testing, the marking and the selection, complicates - or sometimes even obstructs - the use of individual test results for an evaluation of systems.

These three kinds of problems (and there are others) have to be observed in the external evaluation through internal assessment. One clear conclusion is that the testing has to be restricted to subjects or parts of subjects, which make up non-controversy elements, such as basic instrumental skills in reading, writing, mathematics and foreign languages. As can be seen from the previous sections central standardized tests in subjects like history, civics, geography, religion and biology are not used. This means that individual assessment through standardized achievement testing always will be incomplete and therefore always a matter of dispute. And this, in turn, means that a system evaluation on the basis of such a kind of individual assessment has the same kind of incompleteness.

REFERENCES

Grundin, H. Läs- och skrivförmågans utveckling genom skolåren. (Development of Reading and Writing in School). SÖ-FOU 1975:20. Stockholm. 1975.

Marklund, S. Testing e valutazione nelle scuole svedesi. (Testing and Evaluation in the Swedish School). Scuola primaria. Urbino, Italia. 1970.

National Board of Education. Enhetlig skolstruktur och likvärdig utbildningsstandard. (Unified School Structure and Equal Educational Standard). Anslagsäskanden för budgetåret 1979/80 i sammanfattning. Skolöverstyrelsen. Stockholm. 1978.

National Board of Education. Standardprovsresultat i årskurs 6 i sju län. (Results of Standardized Achievement Tests in Grade 6 in Seven Counties). Pedagogiska nämndens verksamhetsberättelse för 1976/77. Skolöverstyrelsen. Stockholm. 1977.

Purves, A.C. Literature Education in Ten Countries. IEA. Stockholm: Almqvist & Wiksell. 1973.